Journal of Neuro-Oncology
https://doi.org/10.1007/s11060-022-04080-x

RESEARCH

A deep learning model for discriminating true progression from pseudoprogression in glioblastoma patients

Mana Moassefi[1,2] · Shahriar Faghani[1,2] · Gian Marco Conte[1,2] · Roman O. Kowalchuk[3] · Sanaz Vahdati[1,2] · David J. Crompton[7] · Carlos Perez-Vega[4] · Ricardo A. Domingo Cabreja[4] · Sujay A. Vora[5] · Alfredo Quiñones-Hinojosa[4] · Ian F. Parney[6] · Daniel M. Trifiletti[7] · Bradley J. Erickson[1,2]

Received: 12 May 2022 / Accepted: 25 June 2022

Abstract

Introduction Glioblastomas (GBMs) are highly aggressive tumors. A common clinical challenge after standard of care treatment is differentiating tumor progression from treatment-related changes, also known as pseudoprogression (PsP). Usually, PsP resolves or stabilizes without further treatment or a course of steroids, whereas true progression (TP) requires more aggressive management. Differentiating PsP from TP will affect the patient's outcome. This study investigated using deep learning to distinguish PsP MRI features from progressive disease.

Method We included GBM patients with a new or increasingly enhancing lesion within the original radiation field. We labeled those who subsequently were stable or improved on imaging and clinically as PsP and those with clinical and imaging deterioration as TP. A subset of subjects underwent a second resection. We labeled these subjects as PsP, or TP based on the histological diagnosis. We coregistered contrast-enhanced T1 MRIs with T2-weighted images for each patient and used them as input to a 3-D Densenet121 model and using five-fold cross-validation to predict TP vs PsP.

Result We included 124 patients who met the criteria, and of those, 63 were PsP and 61 were TP. We trained a deep learning model that achieved 76.4% (range 70–84%, SD 5.122) mean accuracy over the 5 folds, 0.7560 (range 0.6553–0.8535, SD 0.069) mean AUROCC, 88.72% (SD 6.86) mean sensitivity, and 62.05% (SD 9.11) mean specificity.

Conclusion We report the development of a deep learning model that distinguishes PsP from TP in GBM patients treated per the Stupp protocol. Further refinement and external validation are required prior to widespread adoption in clinical practice.

Keywords Glioblastoma · Deep learning · Pseudoprogression · Artificial intelligence · Magnetic resonance imaging

Mana Moassefi and Shahriar Faghani are co-first authors.

✉ Bradley J. Erickson
bje@mayo.edu

1 Artificial Intelligence Laboratory, Mayo Clinic, Rochester, MN, USA

2 Department of Radiology, Mayo Clinic, Rochester, MN, USA

3 Department of Radiation Oncology, Mayo Clinic, Rochester, MN, USA

4 Department of Neurosurgery, Mayo Clinic, Jacksonville, FL, USA

5 Department of Radiation Oncology, Mayo Clinic, Phoenix, AZ, USA

6 Department of Neurologic Surgery, Mayo Clinic, Rochester, MN, USA

7 Department of Radiation Oncology, Mayo Clinic, Jacksonville, FL, USA

Abbreviations

MRI	Magnetic resonance imaging
PsP	Pseudoprogression
TP	True progression
CCRT	Concurrent chemoradiation therapy
DL	Deep learning
RT	Radiotherapy
TMZ	Temozolomide
GBM	Glioblastoma
MGMT	O-6-methylguanine-DNA methyltransferase
IDH	Isocitrate dehydrogenase
CET1	Contrast-enhanced T1
FLAIR	Fluid-attenuated inversion recovery
CNN	Convolutional neural network
RANO	Response assessment in NeuroOncology
ACC	Accuracy
AUROCC	Area under the ROC curve

Published online: 19 July 2022

Introduction

Glioblastoma (GBM), the most common primary brain tumor in adults, has distinctive imaging, molecular and histological features [1]. Despite multimodal treatment, the median overall survival is less than two years [2]. Currently, the most effective treatment for GBM includes maximal surgical resection followed by concurrent chemoradiation therapy (CCRT) with or without tumor-treating fields, and then adjuvant temozolomide (TMZ), known as the Stupp protocol [3]. However, in many cases after the end of radiation treatment, magnetic resonance imaging (MRI) may show new or enlarging lesions that improve without additional intervention besides adjuvant TMZ over time; this phenomenon is known as pseudoprogression (PsP) and is seen in 15% to 36% of GBM patients [4–7].

PsP can mimic disease progression in post-radiation imaging, most commonly 3 to 9 months after radiation [8, 9]. True progression (TP) can also occur in this same time period. MRI is often unable to distinguish these entities. The diagnosis of PsP is typically based on an improvement or stabilization of imaging and clinical findings after a period of continued temozolomide without increased steroids or other potential confounders. Misdiagnosis of PsP and TP can lead to suboptimal clinical care and contribute to patient anxiety [10]. PsP indicates that the current therapy should be continued and is associated with improved survival, even compared with subjects with no new lesions [11–13]. On the other hand, TP indicates that the current treatment is ineffective and needs to be changed [14].

In clinical trials and clinical practice, Response Assessment in NeuroOncology (RANO) criteria are widely used to determine treatment response [15]. The RANO criteria entail evaluating treatment responses based on MRI and clinical information, assuming that MRI measurements of a lesion's largest 2-Dimensional (2D) measures are surrogate markers of tumor size. However, this is not always the case in clinical practice. Brain tumors can display complex shapes and exhibit anisotropic growth which can be influenced by marginal anatomic boundaries, host tissue-tumor interface, and treatment-related effects [16–18]. Studies have suggested that volumetric or 3D measurements have more accurate and reliable assessment outcomes than 2D measurements [18–20]. When evaluating tumor burden and response, volumetric assessment is a crucial parameter.

Studies have shown that deep learning (DL) can be used to classify glioma tumor characteristics and outcomes in a variety of settings [21–23]. In particular, convolutional neural networks (CNNs) have demonstrated excellent performance regarding image recognition and classification [24]. Different features are extracted from the input images through convolution operations across multiple layers in these networks. Based on these learned features, a model can return output such as the absence or presence of the tumor, tumor grades, etc. Several studies have compared the performance of CNNs and other machine learning techniques in medical imaging. Thus, CNNs are becoming more popular among DL techniques for medical imaging because of their low error rate and good accuracy across multiple types of medical images [25, 26].

Using CNN to discriminate PsP from TP has been explored in several articles [27, 28]. In this study, we developed a DL model to distinguish PsP from TP in GBM patients noninvasively using only routine follow-up MRI sequences.

Method

Our institutional review board approved the study and granted a waiver of informed consent. Methods were conducted per regulatory requirements.

Study population

We included histopathologically proven GBM patients according to the 2016 WHO classification who went through maximal surgical resection followed by intensity-modulated adjuvant radiotherapy (RT) with a conventional fraction (2GY per fraction) with or without tumor treating fields and concurrent temozolomide, and after that receiving monthly temozolomide cycles. They all met the inclusion criteria described in Table 1 [29]. We reviewed clinical notes and radiology reports of all cases for one year after surgery and selected cases showing a new enhancing lesion within the original radiation field or an increase in the size of an existing lesion qualifying as progressive disease (PD) [15].

Definition of progressive disease

Patients with GBM who follow the Stupp protocol do not all show progression symptoms. Many remain stable and respond well to treatment. For distinguishing those with progressive disease course, we used RANO criteria response assessment tool. Rather than using the 2D measurement described in RANO, we used a volumetric analysis, so -as we described in the introduction-, instead of including patients with a 25% increase in the largest 2D tumor radius, we included patients with a 40% increase in 3D CE volume [30]. To measure volumes, we used an automated segmentation and volumetric measurement tool: HD-GLIO-AUTO [30–32] (https://github.com/NeuroAI-HD/HD-GLIO-AUTO). HD-GLIO-AUTO is a DL-based tool that is

Table 1 Inclusion and exclusion criteria

Inclusion criteria	
1	Pathologically confirmed glioblastoma. (Primary tumors only–no recurrent cases.)
2	At least 18 years old at initial diagnosis
3	CCRT with TMZ and adjuvant TMZ received after surgical resection or biopsy
4	Availability of CET1, T1, T2, and FLAIR at baseline and at follow-up
Exclusion criteria	
1	Deviation from the standard treatment protocol (e.g., use of second-line treatments like monoclonal antibodies, radiation dose < 30GY), etc
2	No new enhancing lesion; new enhancing lesion outside the original radiation field; new enhancing lesion below our volumetric cut-off
3	Inadequate MRI quality (e.g., movement artifacts)
4	MRI was acquired before 2001
5	History of previous intracranial surgery unrelated to glioblastoma

designed for automated preoperative tumor segmentation. However, we applied it to postoperative images, so we manually checked all the extracted masks and made corrections when necessary. We collected the first images that showed the new lesion in the radiation field or the increase in size based on the radiologist's reports and where the neuroradiologist could not determine whether the change was due to TP or PsP. We reviewed all charts and progress notes for the patients and excluded those who met one of our exclusion criteria (Table 1).

Categorization as TP or PsP

For each patient, we reviewed the following radiology reports and clinical notes for one year after the progression to see if the lesion and patient's clinical status stabilized, improved, or progressed further after the index MRI (the one showing possible progression). We labeled those who remained stable or improved in both imaging and clinical status (without the treatment change) as PsP and those with further imaging and clinical deterioration as TP. The patients for whom a new therapy (e.g., bevacizumab) was initiated were excluded, as it was difficult to determine whether they were TP or PsP-if the improvement was due to the treatment or resolved spontaneously. The final labels of TP and PsP were based on a retrospective review of the diagnoses made at the time of a patient's clinical care. A sub-cohort of subjects underwent a second resection. We labeled these subjects as PsP or TP based on the histological diagnosis. From these, we created a dataset of patients with a suspicious lesion with a strong basis for being labeled as PsP or TP.

Dataset creating and preprocessing steps

We collected fluid-attenuated inversion recovery (FLAIR), T1, (contrast-enhanced T1) CET1, and T2 for the time point when the MRI showed a suspicious lesion. We reviewed all sequences to ensure appropriate image quality.

We used HD-GLIO-AUTO to segment the MRI exams resulting in enhancing lesion volumes. Specifically, the pipeline performed brain extraction, and co-registration of all 4 sequences, and then extracted the tumor masks for the scans and measured the volumes [18, 32]. The pipeline uses HD-BET for brain extraction [31] which demonstrates robust performance in the presence of pathology or treatment-induced tissue alterations and is applicable to a variety of MRI sequence types and does not depend on the MRI vendor types or the acquisition parameters. The brain extracted CET1, FLAIR, and T2 images were co-registered on T1 for each study using the linear image registration tool, FMRIB software library, FSL [33].

Model architecture and training process

We split our dataset at the patient level (to avoid information leakage) into 5 folds stratified by the class (TP and PsP) status using the Scikit-learn library. We applied five-fold cross-validation to assess the model's robustness to perturbations in the data. We tested different combinations of all 4 sequences and found that CET1 and T2 provided us with the best results. We stacked the T2 and CET1 sequences as different channels and normalized the voxel intensity of each channel to zero mean and unit standard deviation. The model took two single-modality MRI sequences (T2, CET1) as input and then concatenated these modalities as a 4-dimensional matrix (2, l, h, w), where "2" represented 2 channels and (l, h, w) represented the shape size (length, height, width) of each single-modality MRI sequence data. We resized the width and height of each channel to 256 × 256

pixels and used symmetric zero padding for the z dimension and padded all the images to the highest z dimension value in our dataset (256). To increase the generalizability of the model and reduce the risk of overfitting, rotation 15°, 15°, and 12° in x, y, and z dimensions respectively with 0.2 probability, flipping in x dimension with 0.5 probability, scaling 0.05 in x, y, and z dimensions, and translation 15,15,5 voxels in x, y, and z dimensions respectively were utilized as augmentation methods.

Due to the relatively small dataset size, we took advantage of transfer learning. We pretrained a 3D-DenseNet-121 [34] classifier on a portion of the BraTS-Reg dataset [35] (https://www.med.upenn.edu/cbica/brats-reg-challenge/) to identify pre versus post-surgical images, and then we used this model's weights and finetuned our main classifier.

We trained a 3-D DenseNet-121 model from the MONAI package [36, 37]. Figure 1 shows a schematic diagram of our training pipeline. DenseNet is a CNN that connects each layer to every other layer in a feed-forward fashion and extracts important features using convolutions (Supplementary Figure-1 in supplementary information). DenseNet introduces a 1×1 convolution as a bottleneck layer before each convolution layer to reduce the number of feature maps. Each dense block contains the bottleneck structure and dimension reduction in transition layers, making parameters more efficient and reducing model complexity. Since our dataset had 124 samples, using DenseNet as building blocks along with transfer learning reduces the risk of overfitting and accelerates training.

We used the Adam optimizer with a mini-batch size of 4. The learning rate was set to 1e-5 for 100 epochs. We trained our model on a cluster of 4 ≈NVIDIA-A100 GPUs.

Given that there was no major class imbalance between the PsP and TP cases, we did not use any sampling method to solve the imbalance. We evaluated our model using overall accuracy (ACC). the area under the ROC curve (AUROCC), sensitivity, and specificity.

For each fold, we trained the model 3 times with different weight initializations and each time the best model's performance was saved. We ensembled these 3 models for each fold and reported that as the final result for that fold.

We experimented with adding age and sex as input features and training a multimodality model, but our results showed no improvement in the prediction and the model performed slower, so we decided not to use this information.

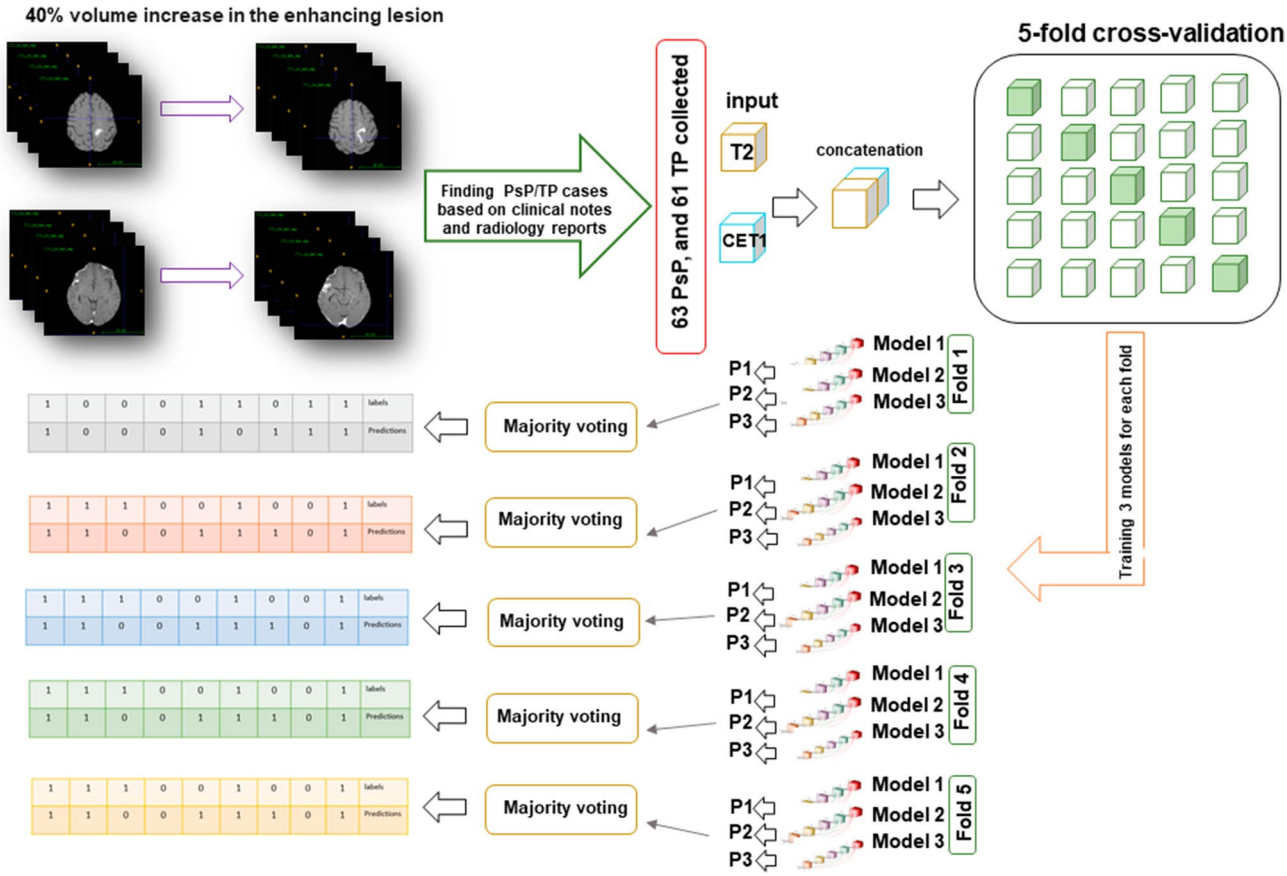

Fig. 1 Brief schema of our prediction pipeline. *PsP* Pseudoprogression, *TP* true progression, *CET1* contrast-enhanced T1, *P* Prediction

Statistical analysis

We performed the Kolmogorov–Smirnov Test of Normality to check the normal distribution of ages in all datasets. In order to compare the differences in clinical characteristics between the PsP and TP groups, we performed Fisher's exact test and chi-square test for categorical variables (sex, O-6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status, Isocitrate dehydrogenase (IDH) mutation status, and RT dose (we classified RT dose to 60cGY and more and less than 60cGY [38]), and an unpaired t-test for continuous variables (age). We used SPSS 20.0 (SPSS Inc., Chicago, IL, USA) software for the analysis. An indication of statistical significance was a P value less than 0.05.

Results

Characteristics of the dataset

A total of 124 cases were collected. Patient characteristics are summarized in Table 2. After the screening process explained in the method Sect. 63 brain lesions (50.80%) were classified as PsP and 61 (49.20%) as TP (Fig. 2). The value of the Kolmogorov–Smirnov test statistic (D) was 0.0857. The p-value was 0.30466 showing that the age distribution of the patients in the dataset did not differ significantly from that which is normally distributed. There were no significant differences in age, gender, IDH mutational status, and RT dose but the MGMT promoter methylation status showed a significant difference between the groups (Table 2).

In predicting PsP versus TP task, our model achieved 76.4% (SD 5.122) mean accuracy, 0.7560 (SD 0.069) mean AUROCC, 88.72% (SD 6.86) mean sensitivity, 62.05% (SD 9.11) mean specificity. The results of each fold are shown in the Table 3.

Discussion

In this study, we developed a DL model which can discriminate PsP from TP with 88.72% sensitivity (over 90% showed in 4 folds), and 76.4% accuracy in GBM patients using multiparametric MRI information where the original neuroradiologist interpretation indicated a high level of uncertainty. We conclude that T2 and CET1 images that can be helpful in predicting a new lesion's growth to be a tumor recurrence or regress and be a therapeutic response. To our knowledge, this is one of the largest datasets of GBM patients classified as PsP and TP, with 124 cases in total.

Several articles have used artificial intelligence-based methods to predict the outcomes of GBM patients and specifically predicting PsP and TP [27, 28, 39–46]. Among AI-based methods, using radiomics and extracting features from that has been a popular technique [39, 45]. The radiomics feature extraction pipeline is completely dependent on a highly accurate tumor segmentation process. To the best of our knowledge, currently, there is no free available automatic tool to perform the postoperative MRI segmentation, so the segmentation needs to be done by neuroradiologists. This process is very time-consuming and expensive. This is a major limitation to the routine application of radiomics-dependant methods.

Table 2 Patient characteristics (*TP* true progression, *PsP* pseudoprogression, *MGMT* O-6-methylguanine-DNA methyltransferase, *IDH* Isocitrate dehydrogenase, *RT* radiotherapy, *GY* Gray)

Dataset	PsP (N = 63)		TP (N = 61)		*P*	Total
	N	%	N	%		
Age (median, range)	59.668 (34–84)		56.94 (24–68)		.36	57.84 (24–84)
Gender					.30	
Male	40	63.49	44	72.13	84 (67.74%)	
Female	23	36.51	17	27.87	40 (32.26%)	
MGMT promoter methylation status					0.00	
Methylated	16	25.49	3	4.91	19 (15.33%)	
Unmethylated	6	9.52	15	24.60	21 (16.93%)	
Unknown	41	65.08	43	70.49	84 (67.74%)	
IDH mutational status					1.00	
Mutated	2	3.17	2	3.27	4 (3.22%)	
Wild type	26	41.27	20	32.79	46 (37.10%)	
Unknown	35	55.56	39	63.94	74 (59.68%)	
RT dose					.76	
60GY or more	46	73.02	46	75.41	92 (74.19%)	
Less than 60 GY	17	26.98	15	24.59	32 (25.81%)	

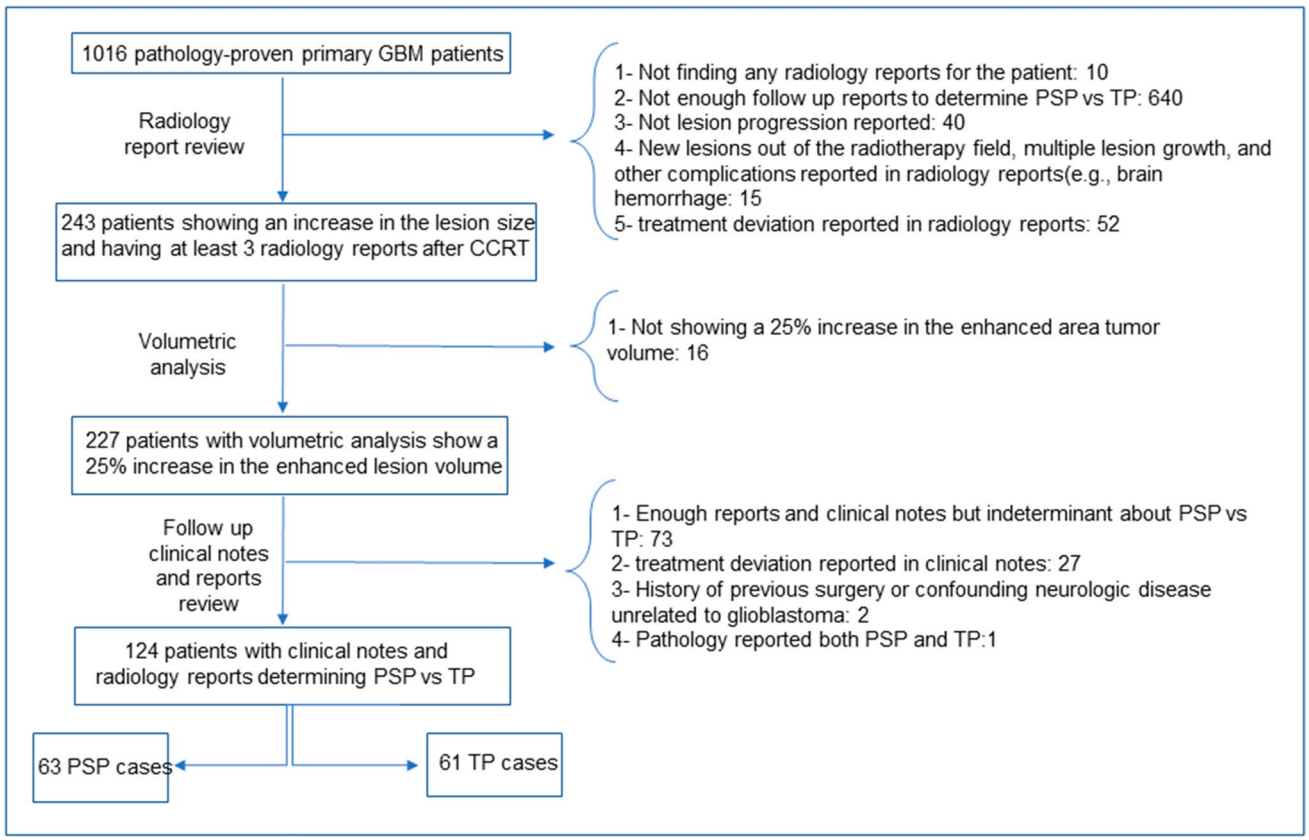

Fig. 2 Flowchart of patients screening. (*GBM* Glioblastoma, *PsP* pseudoprogression, *TP* true progression, *CCRT* concurrent chemoradiation therapy)

Table 3 Outcome of the CNN prediction model in the validation group (n = 25)

	Fold 1 (%)	Fold 2 (%)	Fold 3 (%)	Fold 4 (%)	Fold 5 (%)
Accuracy (first model)	62.00	76.00	80.00	65.00	76.00
Accuracy (second model)	66.00	72.00	80.00	65.00	72.00
Accuracy (third model)	66.00	80.00	80.00	65.00	76.00
Ensemble accuracy result	72.00	80.00	84.00	70.00	76.00
Ensemble AUROCC result	0.72	0.81	0.85	0.66	0.75
Ensemble sensitivity result	92.31	92.31	91.67	75.00	92.31
Ensemble specificity result	50	66.67	76.92	58.33	58.33

Most articles rely on classic machine learning techniques for the classification task [39, 41–43, 45, 46]. There are a few articles that applied DL. One of them is the study in which they applied CNN on the diffusion tensor imaging (DTIs) to extract the imaging features relevant to PsP and TP classification and then radiologists selected the highest relevant features and layers of the images to feed to the next DL classifier [27]. There is another article based on a deep convolutional generative adversarial network (DCGAN), to discriminate between PsP and TTP again in DTI images, then they used the discriminative features for classification by a support vector machine [44]. Another study convolutional applied a recurrent neural network-based DL structure to discriminate between PsP and TP using multi-parametric MRI on 43 GBM patients (7 PsP, and 36 TP cases) [28]. The study concluded that feeding the model with multiple pulse sequences is more efficient than feeding it with only one. We tried to address some of these studies' limitations. One of the greatest advantages of CNNs for image classification is that it eliminates the complicated feature-engineering part of the traditional approach, also no further intervention by the neuroradiologists to predict PsP versus TP. We have used a fivefold cross-validation which kept 24 to 25 cases in each model as the validation set. The validation set of 25 cases in

each fold lowered the probability of the model making the correct predictions at random.

There are several articles showing the attempt to use advanced MR techniques in determining PsP from TP, and also studies using them in conjunction with AI demonstrating variable results [8, 28, 42, 43, 47–49]. We have only used commonly acquired MRI sequences: T1, CET1, T2, and FLAIR. We did a series of experiments with different combinations of these sequences, and we got our best performance using CET1 and T2, so our model only needs these 2 MRI sequences which are available in routine follow-up. Generally, we found our dataset's T2 images to be more consistent and FLAIR images more variant, so that T2 sequence inclusion would allow for more robust image processing. Models which predict the outcome using DTI and perfusion MRI may not be useful in all areas since not all medical facilities have these modalities for the patients' follow-up.

MGMT promoter methylation status and IDH mutation status have proved to be predictive of patient overall survival in GBM patients [50]. In our dataset even though MGMT promoter methylation status showed a difference in PsP and TP patients, we have not used MGMT promoter methylation status and IDH mutation status in our model training because we had the information for less than 50% of our dataset. We acknowledge that GBM diagnosis based on WHO 2020 requires the patients to be IDH wildtype, and we had 2 IDH mutated patients, but since we do not have the IDH information for 63.94% of our cases we did not find it rational to exclude those 2 and included the patients based on WHO 2016 classification.

Clinically, we have accurately recorded the time when we collected the imaging data. It is also crucial to determine if they predict PsP and TP using the baseline image after the surgery or if they are using a progression time image. Unfortunately, none of the articles we reviewed provided such detailed information. In stating the time point, we have been very accurate, and we have only included the controversial images. We have screened the patients based on detailed inclusion and exclusion criteria; for example, new lesions must remain within the original radiation field, otherwise, they are a true progression, and they cannot be considered in the PsP.

In conclusion, advanced computational analyses are increasingly used in the clinical evaluation of GBM and their response to treatment. The present study uses structural MRI modalities to develop an imaging signature that may noninvasively distinguish TP from PsP. Accurate stratification of TP and PsP may facilitate appropriate triage of patients to continue maintenance, therapy, or evaluation for a new intervention, which carries great importance in the era of increasing personalization of therapy. In this paper, we presented our initial investigation of DL in predicting therapy response and outcome in GBM patients. A proof-of-concept study such as this was our first step toward incorporating more information such as treatment details (e.g., bevacizumab) in subsequent studies. One limitation of our study includes the fact that data relied on a pool from a single institution and could benefit from validation in multi-institutional data.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11060-022-04080-x.

Acknowledgements None

Author contributions M.M and S.F wrote the main manuscript text. M.M, S.F, and G.C. trained the model. M.M., S.V., C.P., R.K., A.Q., I.P., S.V., D.C., and R.D. worked in the data collection and dataset gathering process. B.J. and D.T. supervised the project and modified the text.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Tamimi AF, Juweid M (2017) Epidemiology and outcome of glioblastoma. In: De Vleeschouwer S (ed) Glioblastoma. Brisbane, AU: Codon Publications
2. Tesileanu CMS, Dirven L, Wijnenga MMJ, Koekkoek JAF, Vincent AJPE, Dubbink HJ et al (2020) Survival of diffuse astrocytic glioma, IDH1/2 wildtype, with molecular features of glioblastoma, WHO grade IV: a confirmation of the cIMPACT-NOW criteria. Neuro Oncol 22:515–523
3. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB et al (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352:987–996
4. Chamberlain MC, Glantz MJ, Chalmers L, Van Horn A, Sloan AE (2007) Early necrosis following concurrent Temodar and radiotherapy in patients with glioblastoma. J Neurooncol 82:81–83
5. Brandes A, Tosoni A, Franceschi E, Blatt V, Ermani M (2007) Pseudoprogression after concomitant radio-chemotherapy treatment in newly diagnosed glioblastoma patients and potential correlation with MGMT methylation status. Neuro Oncol Oxford University Press (OUP) 9:529–529
6. Chaskis C, Neyns B, Michotte A, De Ridder M, Everaert H (2009) Pseudoprogression after radiotherapy with concurrent temozolomide for high-grade glioma: clinical observations and working recommendations. Surg Neurol 72:423–428
7. Abbasi AW, Westerlaan HE, Holtman GA, Aden KM, van Laar PJ, van der Hoorn A (2018) Incidence of tumour progression and pseudoprogression in high-grade gliomas: a systematic review and meta-analysis. Clin Neuroradiol 28:401–411
8. Hygino da Cruz LC Jr, Rodriguez I, Domingues RC, Gasparetto EL, Sorensen AG (2011) Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. Am J Neuroradiol 32:1978–1985

9. Brandsma D, Stalpers L, Taal W, Sminia P, van den Bent MJ (2008) Clinical features, mechanisms, and management of pseudoprogression in malignant gliomas. Lancet Oncol 9:453–461
10. Fatterpekar GM, Galheigo D, Narayana A, Johnson G, Knopp E (2012) Treatment-related change versus tumor recurrence in high-grade gliomas: a diagnostic conundrum–use of dynamic susceptibility contrast-enhanced (DSC) perfusion MRI. AJR Am J Roentgenol 198:19–26
11. Clarke JL, Chang S (2009) Pseudoprogression and pseudoresponse: challenges in brain tumor imaging. Curr Neurol Neurosci Rep 9:241–246
12. Bronk JK, Guha-Thakurta N, Allen PK, Mahajan A, Grosshans DR, McGovern SL (2018) Analysis of pseudoprogression after proton or photon therapy of 99 patients with low grade and anaplastic glioma. Clin Transl Radiat Oncol 9:30–34
13. Himes BT, Arnett AL, Merrell KW, Gates MJ, Bhargav AG, Raghunathan A et al (2020) Glioblastoma recurrence versus treatment effect in a pathology-documented series. Can J Neurol Sci 47:525–530
14. Poulsen HS, Urup T, Michaelsen SR, Staberg M, Villingshøj M, Lassen U (2014) The impact of bevacizumab treatment on survival and quality of life in newly diagnosed glioblastoma patients. Cancer Manag Res 6:373–387
15. Chukwueke UN, Wen PY (2019) Use of the Response Assessment in Neuro-Oncology (RANO) criteria in clinical trials and clinical practice. CNS Oncol. 8:CNS28
16. Wick W, Gorlia T, Bendszus M, Taphoorn M, Sahm F, Harting I et al (2017) Lomustine and bevacizumab in progressive glioblastoma. N Engl J Med 377:1954–1963
17. Korn RL, Crowley JJ (2013) Overview: progression-free survival as an endpoint in clinical trials with solid tumors. Clin Cancer Res 19:2607–2612
18. Kickingereder P, Isensee F, Tursunova I, Petersen J, Neuberger U, Bonekamp D et al (2019) Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. Lancet Oncol. https://doi.org/10.1016/s1470-2045(19)30098-1
19. Chow DS, Qi J, Guo X, Miloushev VZ, Iwamoto FM, Bruce JN et al (2014) Semiautomated volumetric measurement on postcontrast MR imaging for analysis of recurrent and residual disease in glioblastoma multiforme. AJNR Am J Neuroradiol 35:498–503
20. Sorensen AG, Patel S, Harmath C, Bridges S, Synnott J, Sievers A et al (2001) Comparison of diameter and perimeter methods for tumor volume calculation. J Clin Oncol 19:551–557
21. Lotan E, Jain R, Razavian N, Fatterpekar GM, Lui YW (2019) State of the art: machine learning applications in glioma imaging. AJR Am J Roentgenol 212:26–37
22. Chang P, Grinband J, Weinberg BD, Bardis M, Khy M, Cadena G et al (2018) Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. AJNR Am J Neuroradiol 39:1201–1207
23. Lao J, Chen Y, Li Z-C, Li Q, Zhang J, Liu J et al (2017) A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. Sci Rep 7:10353
24. CS231n Convolutional neural networks for visual recognition. http://cs231n.stanford.edu/. Accessed 28 Feb 2022
25. Suzuki K (2017) Overview of deep learning in medical imaging. Radiol Phys Technol 10:257–273
26. Truong AH, Sharmanska V, Limbäck-Stanic C, Grech-Sollars M (2020) Optimization of deep learning methods for visualization of tumor heterogeneity and brain tumor grading through digital pathology. Neurooncol Adv. 2:vdaa110
27. Liu X, Zhou X, Qian X (2020) Transparency-guided ensemble convolutional neural network for the stratification between pseudoprogression and true progression of glioblastoma multiform in MRI. J Vis Commun Image Represent 72:102880
28. Lee J, Wang N, Turk S, Mohammed S, Lobo R, Kim J et al (2020) Discriminating pseudoprogression and true progression in diffuse infiltrating glioma using multi-parametric MRI data through deep learning. Sci Rep 10:20331
29. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK et al (2016) The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol 131:803–820
30. Kickingereder P, Isensee F, Tursunova I, Petersen J, Neuberger U, Bonekamp D et al (2019) Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. Lancet Oncol 20:728–740
31. Isensee F, Schell M, Pflueger I, Brugnara G, Bonekamp D, Neuberger U et al (2019) Automated brain extraction of multisequence MRI using artificial neural networks. Hum Brain Mapp 40:4952–4964
32. Isensee F, Jäger PF, Kohl SAA, Petersen J, Maier-Hein KH (2019) automated design of deep learning methods for biomedical image segmentation. http://arxiv.org/abs/1904.08128
33. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62:782–790
34. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 2261–2269
35. Baheti B, Waldmannstetter D, Chakrabarty S, Akbari H, Bilello M, Wiestler B, et al (2021) The brain tumor sequence registration challenge: establishing correspondence between pre-operative and follow-up MRI scans of diffuse glioma patients. http://arxiv.org/abs/2112.06979
36. MONAI (2022) Medical Open Network for AI. https://zenodo.org/record/6114127/export/csl
37. The MONAI Consortium (2020) Project MONAI. https://zenodo.org/record/4323059
38. Fernandes C, Costa A, Osório L, Lago RC, Linhares P, Carvalho B et al (2017) Current standards of care in glioblastoma therapy. In: De Vleeschouwer S (ed) Glioblastoma. Codon Publications, Brisbane (AU)
39. Patel M, Zhan J, Natarajan K, Flintham R, Davies N, Sanghera P et al (2021) Machine learning-based radiomic evaluation of treatment response prediction in glioblastoma. Clin Radiol 76:628.e17-628.e27
40. Akbari H, Rathore S, Bakas S, Nasrallah MP, Shukla G, Mamourian E et al (2020) Histopathology-validated machine learning radiographic biomarker for noninvasive discrimination between true progression and pseudo-progression in glioblastoma. Cancer 126:2625–2636
41. Jang B-S, Park AJ, Jeon SH, Kim IH, Lim DH, Park S-H et al (2020) Machine learning model to predict pseudoprogression versus progression in glioblastoma using MRI: a multi-institutional study (KROG 18–07). Cancers. https://doi.org/10.3390/cancers12092706
42. Kebir S, Schmidt T, Weber M, Lazaridis L, Galldiks N, Langen K-J et al (2020) A preliminary study on machine learning-based evaluation of static and dynamic FET-PET for the detection of pseudoprogression in patients with IDH-wildtype glioblastoma. Cancers. https://doi.org/10.3390/cancers12113080
43. Qian X, Tan H, Zhang J, Zhao W, Chan MD, Zhou X (2016) Stratification of pseudoprogression and true progression of glioblastoma multiform based on longitudinal diffusion tensor imaging without segmentation. Med Phys 43:5889
44. Li M, Tang H, Chan MD, Zhou X, Qian X (2020) DC-AL GAN: Pseudoprogression and true tumor progression of glioblastoma multiform image classification based on DCGAN and AlexNet. Med Phys 47:1139–1150

45. Sun Y-Z, Yan L-F, Han Y, Nan H-Y, Xiao G, Tian Q et al (2021) Differentiation of pseudoprogression from true progressionin glioblastoma patients after standard treatment: a machine learning strategy combinedwith radiomics features from T1-weighted contrast-enhanced imaging. BMC Med Imaging 21:17
46. Jang B-S, Jeon SH, Kim IH, Kim IA (2018) Prediction of pseudoprogression versus progression using machine learning algorithm in glioblastoma. Sci Rep 8:12516
47. Verma N, Cowperthwaite MC, Burnett MG, Markey MK (2013) Differentiating tumor recurrence from treatment necrosis: a review of neuro-oncologic imaging strategies. Neuro Oncol 15:515–534
48. Sundgren PC (2009) MR spectroscopy in radiation injury. AJNR Am J Neuroradiol 30:1469–1476
49. Akbari H, Rathore S, Bakas S, Nasrallah M, Rozycki M, Mohan S et al (2018) NIMG-70. Quantitative image analysis and machine learning techniques for distinguishing true progression from pseudoprogression in patients with glioblastoma. Neuro Oncol 20:191–192
50. Binabaj MM, Bahrami A, ShahidSales S, Joodi M, Joudi Mashhad M, Hassanian SM et al (2018) The prognostic value of MGMT promoter methylation in glioblastoma: a meta-analysis of clinical trials. J Cell Physiol 233:378–386